

COHORT PROFILE

Cohort Profile: The China Health and Nutrition Survey—monitoring and understanding socio-economic and health change in China, 1989–2011

Barry M Popkin,^{1*} Shufa Du,¹ Fengying Zhai² and Bing Zhang²

Accepted 24 September 2009

How did the study come about?

This study began with the desire of the University of North Carolina (UNC), Chapel Hill (CH) Principal Investigator (PI) to be able to examine across space and time the ways economic and social change affected a range of health behaviours in a large country. China was selected because of its unique opening up of its economic and social system. In 1986 no longitudinal surveys existed in China and all surveys were either very narrow health, economic or demographic surveys. Furthermore, no raw data from any survey had been allowed out of the country. Since China's reform and open policy, the country was being transformed from one facing famine and extreme food shortages to one where the food supply addressed basic needs and the initial states of a major transformation of the food distribution and marketing system was occurring. With this as background, two UNC faculty members, Barry Popkin and Gail Henderson, a China scholar, began the process in 1986 of meeting a range of University and government officials to discuss this work.

The Chinese Center for Disease Control and Prevention (CCDC) (formerly the Chinese Academy of Preventive Medicine, CAPM) was selected. The CCDC, under the leadership of Madame Chunming Chen, was very enthusiastic about the challenges of such an approach to understanding the interplay of socio-economic change and health and decided to

invest the academy's own funds to collaborate on this initiative.

An agreement that focused on publication, data ownership and data dissemination of the results was established. No funding agencies thought such an initiative could be undertaken and the data base obtained by the US institution for use and dissemination. As a result, a small grant of \$60 000 from the Carolina Population Center (CPC) was the only available funding to initiate the first survey.

The China Health and Nutrition Survey (CHNS) was established. The goal was to develop a multipurpose longitudinal survey that would allow the group to examine a series of economic, sociological, demographic and health questions of interest to the CAPM and these scholars. Professor John Akin joined Popkin and Henderson in work on the initial effort. Once the group proved able to collect the data and to get the raw data files out of China, a program project application was submitted to National Institute of Child Health and Human Development (NICHD) that included two additional waves of CHNS data collection. This was funded and all subsequent survey waves were funded by NICHD-funded R01s. Only later did the Chinese Ministry of Health decide the surveys were important, and the ministry instructed the collaborating provinces to place priority on long-term collaboration. The active involvement of the Minister of Health and others in the ministry has provided important political protection for the CHNS and also has provided continued counterpart partial financial support.

Chinese colleagues are full and active partners in the CHNS project and have primary responsibility for data collection. Over the life of the project they have spent considerable time in CH working with the UNC–CH group on data analysis and training activities, and they plan to continue these efforts. Professors Keyou

¹ Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA.

² National Institute of Nutrition and Food Safety, Chinese Center for Disease Control and Prevention, Beijing, China.

* Corresponding author. Carolina Population Center, University of North Carolina, 123 W. Franklin Street, Chapel Hill, NC 27516, USA. E-mail: popkin@unc.edu

Ge, then Fengying Zhai, and now Bing Zhang lead the project for the CCDC. Professor Shuigao Jin, a statistician and computer expert, initially headed the component dealing with sampling and database development. Professor Huijun Wang is a key member of field work, data cleaning and data analysis. At one time Professor Shufa Du headed the project in China and now, as a faculty member at UNC, plays a major role in working with the CHNS.

Part of the reason for the careful selection of the collaborator for the CHNS was the desire to insulate this study from the political and institutional problems that have hampered many subsequent survey research projects conducted jointly by US and Chinese institutions. By partnering as we have, we have been successful in continuing this survey during periods in China when many other western-funded social science-related surveys were discontinued.

In 2009 the Department of Laboratory Medicine of China-Japan Friendship Hospital (CJFH), Ministry of Health (MOH) joined this project as the lead agency for the collection, storage and analysis of biospecimens. Dr Shengkai Yan leads the work related to biospecimens.

What does it cover?

This survey was designed to cover key public health risk factors and health outcomes, demographic, social and economic factors in depth at the individual, household and community levels. This in-depth survey collects data on occupations; incomes and benefits of working-age household members; time use; diet and nutritional status; activities of daily living, health status and use of health services; marriages, birth preference and pregnancies of reproductive-age women; mass media and body images; household size and composition; living arrangements; care of children and elders; housing conditions; land ownership; and limited household asset ownership. Individual health-related data are highly detailed, and include carefully measured dietary intake, physical activity, smoking and drinking data, anthropometrics, blood pressure and limited clinical data from all respondents. Beginning with the CHNS 2009, spatial coordinates on all respondents and key community resources, fasting blood for all respondents ≥ 7 years old, and toenail samples for all respondents ≥ 2 years old are being collected. It also collects detailed community economic, social, demographic and infrastructural data.

Who is in the sample?

The CHNS sample was not designed to be representative of China but to be randomly selected and to capture a range of economic and demographic circumstances and to provide data from randomly

selected households in eight provinces—Liaoning, Shandong, Henan, Jiangsu, Hubei, Hunan, Guizhou and Guangxi (from north to south). Data to create a representative provincial level sample could not be obtained in 1988. A multistage, random cluster process was used to draw the sample in each of the provinces. Counties and cities in each province were stratified by income (low, middle and high) and a weighted sampling scheme was used to randomly select four counties and two cities in each province. Villages and townships within the counties and urban and suburban neighbourhoods within the cities were selected randomly. In each community, 20 households were randomly selected and all household members were interviewed; only preschoolers and young adults aged 20–45 years were surveyed in 1989 due to constraints of funding. The current sample consists of 216 communities from nine provinces (Heilongjiang province was enrolled as a ninth province in 1997), comprising of 36 urban neighbourhoods, 36 suburban neighbourhoods, 36 towns and 108 villages. The household sample was 4020 in 1989 and 4467 in 2006. For individuals, it was 15927 in 1989 and 18 764 in 2006.

How often have they been followed up?

The CHNS rounds have been completed in 1989, 1991, 1993, 1997, 2000, 2004, 2006 and now 2009. The CHNS 2011 is also funded and will be undertaken. We will then propose two new rounds to the National Institutes of Health (NIH) with a request for a second round of fasting blood and toenail samples. Due to cancellation of travel linked with the SARS outbreak, the CHNS 2003 was shifted to 2004. All funding has come from the National Institute of Child Health and Development with one small National Science Foundation, an institutional development and training grant from the NIH Fogarty Institute, CPC support for the first wave, and some small foundation support.

The CCDC and provincial CDCs have provided strong cofunding support for local field survey costs, including car, manpower and per diems in the initial four to five surveys.

What has been measured?

Essentially all rounds of the CHNS have collected identical data from the community and household. The CHNS 1989, undertaken as the first survey to collect individual dietary intake data in China, obtained these dietary, clinical and anthropometric data from children aged < 6 years and all adults aged 20–45 years. All subsequent surveys have obtained clinical, dietary, anthropometric and all other individual data from each household member (Table 1).

Table 1 Key domains of data—CHNS

Data domains	CHNS
Samples	Nine provinces
Number of surveys	Seven (1989–2006) ongoing 2009 and 2011 funded
Households individuals (100%)	4600/19 000
Detailed HH composition/education	All surveys
Marriage history	All surveys
Family planning use/birth history	All surveys
Detailed income/employment by sector of work	All surveys
Assets	All surveys
Inter- and intra-generation transfers	All surveys
Time allocation	All surveys
Aging-IADL-ADL	Yes—not all rounds
Detailed health service use/insurance	All surveys
Detailed smoking/drinking	All surveys
Detailed diet/physical activity or inactivity	3 days, weighed and measured
Anthropometry	All surveys
Clinical exams with blood pressure data	All surveys
Hb/blood spots	2009
Fasting blood	2009
Toenails	2009
DNA—blood stored	2009
GPS coordinates	Begin 2009
In-depth contextual data	All surveys

HH: Household; IADL: Instrumental activities of daily living; ADL: Activities of Daily Living; GPS: Global positioning system.

What is attrition like?

Response rates and attrition are very complex to determine with this survey, because the participants who left in one survey year may have moved back in a later year, and because, since 1997, we have recruited new participants as replenishment samples if a community has less than 20 households or if participants have formed a new household or separated from their family into a new housing unit in the same community (Tables 2 and 3). In 1997, one province was unable to participate in the survey for natural disaster, political and administrative reasons, namely the Liaoning Province, which could not take on Beijing-initiated projects. We added a new province, Heilongjiang, in CHNS 1997 and then surveyed both the old returning province and the new one in CHNS 2000 and all subsequent surveys. If we define response rate based on those who participated in previous survey rounds remaining in the current survey,

our response rates were ~88% at individual level and 90% at household level. If we define response rate based on those who participated in 1989 and remained in the last round in 2006, they were 63 and 69%, respectively.

We have had several major causes of loss to follow-up. The first is missing people, whom we could not find, due to travel, hours of work or play, and/or refusal to come for anthropometric and clinical exams. The second is school children who were in boarding schools, a practice that greatly accelerated in 2004 and 2006, and who entered colleges and universities. The third is migrant work for those aged ≥ 16 years. And last, natural disasters and major redevelopment of housing in all large urban centres added to these relocations. For example, 1997 was the year our provinces experienced a major natural crisis. Flooding affected ~25% of our rural sample. The villages were flooded for 3 months and villagers moved away in scattered fashion, but most individuals returned later. In CHNS 2006, we found 13 households in one community of Hunan province and many more in other provinces that had returned the previous year. A new province—Heilongjiang—was added and Liaoning Province rejoined in 2000; therefore, the total sample size increased to approximately 19 000 individuals who lived in 4500 households.

Table 3 shows that >75% of households participated in five to seven surveys and an additional 17% in three to four surveys. We defined individual participation very conservatively as those who provided 3 days of individual dietary data, employment, occupational and income data (if an adult). First, as noted, 7.25% of the sample has died; 66% have been in four or more rounds of the survey (Table 3). These levels of individual response will be increased significantly with the survey of missed individuals during the Chinese New Year proposed for the new rounds. Details on the survey in general are on the CHNS web site: <http://www.cpc.unc.edu/projects/china>.

Our lost-to-follow-up rate increased in the past two waves. To remedy these problems, our team, along with provincial collaborators, spent >18 months discussing options. We have piloted new strategies that will be instituted in 2009 and refined for 2011. The first change is to have interviewers return to the communities during the Spring Festival to attempt to locate families and individuals we previously missed. The second change is to go to the boarding schools of all children during the week and obtain their dietary, anthropometric and other relevant data that parents will not be able to report (e.g. television viewing, physical activity). This second change seems to work and will be used as one option for both urban and rural areas to attempt to increase the age 4–17 population response rate to 90%. We did some research in two provinces around the Spring Festival re-interview but could not get an adequate sample from which to prepare exact statistics. We expect to see a large

Table 2 CHNS response rates at the individual and household levels

	1989	1991	1993	1997	2000	2004	2006
Individual							
N	15 927	14 789	13 900	15 874	17 054	16 129	18 764
Response rate (%) ^a	100	88.1	88.2	80.9	83.0	80.2	88.0
Response rate (%) ^b	100	88.1	80.1	65.8	64.1	52.1	62.9
Household							
N	4020	3819	3621	4022	4515	4416	4467
Response rate (%) ^a	100	94.9	93.5	79.3	89.4	85.9	89.9
Response rate (%) ^b	100	94.9	90.0	72.1	77.2	68.4	68.5

^aBased on previous year.^bBased on 1989 samples.**Table 3** CHNS survey participation rate for 1989 households and individuals

Rounds participated	Household		Individual	
	Percentage	Cumulative percentage	Percentage	Cumulative percentage
All seven rounds	54.03	54.03	24.28	24.28
Six rounds	13.23	67.26	12.99	37.27
Five rounds	8.55	75.81	17.85	55.12
Four rounds	6.96	82.77	10.86	65.98
Three rounds	9.98	92.75	15.78	81.76
Two rounds	4.63	97.38	9.56	91.32
One round	2.62	100.00	8.98	100.00

Table 4 Percent of 1989 participants who died between the prior survey and 2006

1991	1993	1997	2000	2004	2006	Total
1.16%	1.01%	1.69%	1.23%	1.59%	0.57%	7.25%

increase because many migrant farmers who work in other areas will return to their home during the Spring Festival. In four provinces we also tested undertaking the survey in all township boarding schools and this captured most of the children missing from the household surveys for CHNS 2006. Both of these changes will be implemented for all provinces in the CHNS 2009.

Table 4 shows the percentages of the study participants who have died. Because the number of deaths is substantial, analyses have been published with these individuals as a focus.

What has it found? Key findings and publications

The CHNS has been used for an array of research across dozens of disciplines and journals. The impact

can be measured by the publications of the individuals and institutions that have downloaded the data. Among important policy results have been research used to launch a national fortification law for fortifying soy sauce and flour with iron and other micronutrients, the basis for World Bank poverty reduction programs in a number of poor counties, a large series of studies to justify national programs linking agricultural price policy with nutrition to attempt to increase soybean consumption and other price research being used to consider some ways to reduce caloric intake and weight gain, and the basis for many national and provincial programs addressing poor dietary and activity patterns and increased obesity.¹⁻³ The obesity work has shown how shifting from bikes and walking to cars and motorcycles has increased incident obesity, and how declines in home production and shifts toward more reduced-activity market-related occupations have increased the risk of obesity.^{4,5} Other work in nutrition and chronic disease has shown the body mass index-hypertension relationships in Chinese people are much steeper than in the USA among Blacks and Whites,⁶ and the shifting dietary and activity patterns among the poor that have heralded an increased burden of obesity among the poor.⁷

Table 5 Master longitudinal CHNS files

Data files	Brief description	Number of variables
Two identification files	Identification (gender, birth date, interview date, death date)	100
Two physical education and physical activity files	Physical exam and physical activity/inactivity	594
Eight diet and infant feeding files	Household level, individual meal level, individual level food files, and infant feeding practice	164
20 income-related files	Education, occupation, wage, household assets	635
Nine child care and ever-married women files	Child care activity and availability, child bearing, mass media, marriage, pregnancy, and birth history	342
Three health service files	Medical insurance, health care and availability	136
Seven community files	Infrastructure, family planning, food pricing, medical services, health facilities, television channels and urbanization index	1751
Total	51 files including 7 community datasets	3722

What are the main strengths and weaknesses?

The major strength of the CHNS is the ability to capture enormous heterogeneity and change spatially and temporally in one of the most rapidly changing environments in the world. Because of its long duration and wide geographic coverage, the CHNS can document the dramatic economic, social, behavioural and health status changes that have characterized China in the past several decades. Key weaknesses include the loss-to-follow-up data, particularly among individuals. This came partially because of the PI's cautiousness in obtaining adequate NIH funding. A second major weakness is the lack of sampling weights, an issue that we wish we had tried to address in 1987 through more efforts to work with the State Statistical Bureau and involvement of a more experienced Chinese sampling researcher. In addition, we would have liked to collect GPS data earlier, but the Chinese government did not allow this until 2009 and we would have liked to have foreseen more of the major food system changes and obtained contextual data to capture the early stages of these changes (mainly the shift toward modern large supermarket food systems and more data on media exposure).

Can I get hold of the data? Where can I find out more?

All data are available free to scholars at the web site <http://www.cpc.unc.edu/projects/china/data/data.html>. We have created 51 new datasets, known as CHNS Longitudinal Master Files, available for download in SAS transport format, including those shown in Table 5. These new master files are designed to make longitudinal analysis of the CHNS Survey data much easier. All variable names are standardized across all

survey years when they represent the same survey question. With a few exceptions, they are standardized to the 3700 variable names. Data formatting is standardized across all survey years.

The exceptions are requests for information that can provide ways to locate the sample. The contextual data require some confidentiality forms be completed prior to release. The Food Composition Table data are not available as they are controlled by separate Chinese authorities. GPS coordinates will not be provided; however, procedures to pay for linkages and for collection of additional coordinates will be provided on the website. Also, by the middle or end of 2011, provisions for obtaining access to buffy coat, biomarker and other data will be available. All biomarker data are stored at the CJFH.

Funding

National Institutes of Health (RO1-HD30880, DK056350 and RO1-HD38700) and the Fogarty International Center.

Acknowledgements

The authors thank the National Institute of Nutrition and Food Safety, China Center for Disease Control, the National Institutes of Health (NIH) (RO1-HD30880, DK056350 and RO1-HD38700) and the Fogarty International Center, NIH, for financial support for the CHNS data collection and analysis files from 1989 to 2006 and both parties plus the China-Japan Friendship Hospital, Ministry of Health for support for CHNS 2009 and future surveys. They also wish to thank Ms Frances L. Dancy for administrative assistance; Mr James Terry and Dr Phil Bardsley for managing the CHNS file development; our numerous colleagues at UNC and across the USA who have assisted us; and our hundreds of staff in China for

undertaking repeatedly this very complex survey. None of the authors has conflict of interests of any type with respect to this manuscript.

Conflict of interest: None declared.

References

- ¹ Popkin BM, Paeratakul S, Ge K, Zhai F. Body weight patterns among the Chinese: results from the 1989 and 1991 China Health and Nutrition Surveys. *Am J Public Health* 1995;**85**:690–94.
- ² Guo X, Mroz T, Popkin B, Zhai F. Structural change in the impact of income on food consumption in China, 1989–1993. *Econ Dev Cult Change* 2000;**48**:737–60.
- ³ Guo X, Popkin B, Mroz T, Zhai F. Food price policy can favorably alter macronutrient intake in China. *J Nutr* 1999;**129**:994–1001.
- ⁴ Bell A, Ge K, Popkin B. Weight gain and its predictors in Chinese adults. *Int J Obes Relat Metab Disord* 2001;**25**:1079–86.
- ⁵ Bell A, Ge K, Popkin B. The road to obesity or the path to prevention: motorized transportation and obesity in China. *Obes Res* 2002;**10**:277–83.
- ⁶ Bell A, Adair LS, Popkin BM. Ethnic differences in the association between body mass index and hypertension. *Am J Epidemiol* 2002;**155**:346–53.
- ⁷ Du S, Mroz T, Zhai F, Popkin B. Rapid income growth adversely affects diet quality in China – particularly for the poor! *Soc Sci Med* 2004;**59**:1505–15.